



**VIIth Surana & Surana And Shastra IIT Madras,
Intellectual Property Law Moot Competition 2023-2024**

05th - 07th January 2024

VENUE

**Indian Institute of Technology, Madras Campus
Chennai – 600 036**

IN THE HIGH COURT OF HILED

SAMAY SINHA, PLAINTIFF

V.

EPIONA PRIVATE LIMITED, DEFENDANT

Samay Sinha is a luminary in the literary sphere and commenced his literary voyage with the historical novel 'Demelops' in 1999. This seminal creation garnered widespread acclaim, setting the stage for an expansive and influential literary career. Subsequent to the initial success, Samay Sinha continued to produce an extensive body of work that resonated with readers across diverse demographics. His oeuvre spans various genres, including but not limited to historical fiction, each manifestation a testament to his literary dexterity. Samay Sinha has been the recipient of numerous accolades, including the Sahitya Akademi Award, which serve as tangible acknowledgments of his unparalleled contributions to the literary landscape.

EPIONA, a developer and seller of artificial-intelligence software products, creates large language models (LLMs) such as Interactive Creative Pre-Conditioned Metamorphoser (IntCPM), including CPM-1, CPM-2, CPM-3, and CPM-4. LLMs are not traditional software programs; rather, they are "trained" by copying extensive text from various sources into a training dataset. During training, the LLM extracts expressive information from each piece of text, adjusting its output to closely resemble the sequences of words copied from the training dataset.



EPIONA, in its development of IntCPM, utilized training datasets comprising copyrighted works, including books written by Samay Sinha.

On 16th November 2023 Samay Sinha issued a Cease and Desist Notice to EPIONA stating that it utilized various training datasets, with a substantial portion sourced from copyrighted works, including those of Samay Sinha, without due authorization, credit, or compensation.

Notably, he claimed that EPIONA trained Int CPM on BookCorpus, a collection of over 7,000 unique unpublished books, sourced controversially from a website called Smashwords, without proper consent, credit, or compensation to the authors. EPIONA continued this practice with CPM-3, incorporating 15% of the training dataset from two internet-based books corpora, denoted as "Books1" and "Books2." The contents of these datasets remain undisclosed by EPIONA.

The controversial nature of EPIONA's practices extends to the Books2 dataset, estimated to contain about 294,000 titles, sourced from notorious "shadow library" websites such as Library Genesis (LibGen), Z-Library (Bok), Sci-Hub, and Bibliotik, known for aggregating books illegally and making them available through torrent systems.

EPIONA's IntCPM allows users to input text prompts, generating natural language responses. Whether accessed through a web interface or API, IntCPM's functionality relies on the data upon which it is trained. Samay Sinha contended that IntCPM, when prompted to summarize books authored by him, produced accurate summaries based on the allegedly unauthorized ingestion of his copyrighted materials. Samay Sinha stated that these outputs by IntCPM were derivative works or adaptations of his copyrighted books and that at no point did IntCPM reproduce any copyright management information included in their published works.

EPIONA issued a Reply Legal Notice on 14th December 2023 to Samay Sinha's Cease and Desist Legal Notice stating that copyright safeguards the specific manner in which an author conveys an idea, excluding protection for the underlying concept itself, factual details encompassed within the author's articulated communication, or other fundamental elements of creative expression. Consequently, every concept, theory, or fact presented in a copyrighted work becomes promptly accessible for public utilization upon publication. Hence, although an author may register a copyright for their literary work, statistical information related to word frequencies, syntactic patterns, and thematic markers within said work falls beyond the ambit of copyright protection.

The Reply Legal Notice also stated that a tool extracting information about the original work does not constitute infringement, as it does not replicate protected expression. This principle extends to facts presented in the work, as well as its overarching plot or themes. These principles are integral to the overarching objective of copyright law, fostering intellectual and artistic progress, achievable only through the ability of each author to build upon the contributions of others.

The Reply Legal Notice further stated that there is no substantial similarity and to assess the two contested works by (1) disregarding elements that are not eligible for protection, and (2) evaluating whether the remaining protectible elements, in isolation, exhibit substantial similarity. Since IntCPM's creation does not display substantial similarity to the original, it does not qualify as a "copy" or a "derivative work."



EPIONA also stated that its use of copyrighted material amounted to fair use and that fair use is not an infringement.

On 4th January 2023, Samay Sinha filed this lawsuit in response to EPIONA's unauthorized use of his registered copyrighted materials as training material for IntCPM, without obtaining consent, providing credit, or offering compensation, thereby violating the Indian Copyright Act, 1957. The suit claims:

- A. Interim injunctive relief prevent further unauthorized use of Plaintiffs' copyrighted materials;
- B. An award of damages for the copyright infringement;
- C. Final Injunctive relief to prevent further unauthorized use of Plaintiffs' copyrighted materials;
- D. Accounts of profits derived from the unauthorized use of Plaintiffs' copyrighted materials;
- E. Attorneys' fees and costs incurred in bringing this action; and
- F. Any further relief the Court deems just and proper.

EPIONA is incorporated in Chennai, State of Tamil Nadu. The High Court of Hiled (with the same jurisdictional powers as the Hon'ble High Court of Delhi) is claimed as the proper judicial venue and having jurisdiction over the subject matter of this action as the alleged acts of copyright infringement and related claims fall within its territorial jurisdiction, as the Defendant engages in online activities within this jurisdiction, and a substantial part of the events giving rise to the claims occurred herein. Plaintiff resides in Kolkata, State of West Bengal.